Identification and study of white dwarfs in J-PLUS Data Release 1

Author: A. Cubero
Tutor: J.M. Carrasco

ABSTRACT

Context. The Javalambre Photometric Local Universe Survey, with 12 wide and narrow photometric bands, enable the reconstruction of spectral information that can be key to characterize the astronomical sources observed in the survey.
Aims. Making use of the data published in the first data release we build a method to identify white dwarfs using photometry and parallaxes, to obtain a catalogue of WDs candidates.
Methods. The classification is made using the random forest classifier, a machine learning algorithm. The first step is to select available sources to build training and test datasets. For this we use the available white dwarfs catalogue using information from *Gaia* DR2 and select the rest of objects (non-WDs) using sources outside the white dwarfs locus in the color-magnitude diagram. Different attempts of the classification model are done and discussed, until finding the best option incorporating progressively the information from the 12 J-PLUS passbands, 3 *Gaia* passbands, parallaxes from *Gaia* DR2 and their errors.
Results. Parallax is the most important information for the classification of WDs, as well as the extreme bands, the bluest (u, J0378, J0395) and the reddest (i, J0861, z). The resulting candidates are well placed in many diagrams. The final catalogue contains 24634 white dwarf candidates present in J-PLUS first data release, with information from 15 different filters (12 from J-PLUS and 3 from *Gaia*), parallax and a probability evaluating how reliable the classification is.
Conclusions. The classification model developed is be a very promising tool for the identification of white dwarfs using J-PLUS data, which will be improved in future data releases using data for all the northern celestial hemisphere.

Key words. catalogue – white dwarfs

Contents

1. Introduction

White dwarfs (WDs hereafter) represent the most common end of the evolution of stars in a large range of masses, between 0.07 and 8 $M_{\odot}$ (Ibeling, Heger 2013; Doherty et al 2015; Woosley, Heger 2015; Lauffer, Romero, Kepler 2018). The evolution of stars mainly depends on their mass and metallicity, but ultimately all stars end up collapsing due to the exhaustion of their nuclear fuel. For the vast majority of stars (over 97%), when Helium in the core is exhausted, the star becomes unstable and experiences thermal pulses ejecting the outer envelopes, leading to a carbon core star that slowly starts to cool, becoming a WD. The broad mass range of stars that end up as WDs, make them one of the most interesting populations in order to study stellar evolution (Torres et al. 2019; Tremblay et al. 2019), open and globular clusters (García-Berro et al. 2010; von Hippel, & Gilmore 2000; Hansen et al. 2013; Torres et al. 2015; Campos et al. 2016), structure and kinematics of galactic disks (Chiba, & Beers 2000; Rowell, & Kilic 2019; Reid 2005) or the origin and evolution of the Milky Way and the Universe (Isern et al. 2013; G. Fontaine, P. Brassard, and P. Bergeron 2001; Winget et al. 1987; Garcia-Berro et al. 1988).

Getting the ability to study such amount of aspects depends on the quality and completeness of available WDs catalogues. Due to their low intrinsic luminosities, those complete WDs samples have been a challenge for many years. More recently, the availability of large-scale surveys is providing a great amount of information enabling a better study of the different stellar populations. There are lots of examples of these large-scale surveys, such as the Sloan Digital Sky Survey (York et al. 2000), the LAMOST (Zhao et al. 2012), the PanSTARRS collabora-

tion (Tonry et al. 2012), J-PLUS (Cenarro et al. 2019), and the *Gaia* mission (Gaia Collaboration et al. 2016). The *Gaia* mission has already provided unprecedented positional and radial velocity measurements with great accuracies, producing a census of more than one billion sources. The second data release (*Gaia*-DR2, Gaia Collaboration et al. 2018), allowed the progress in many fields, including the study of the WDs population, the identification of extremely low-mass WD candidates (Pelisoli, & Vos 2019) and the identification of field WDs in the absolute magnitude versus colour diagram, leading to the selection of ~260.000 high-confidence WDs candidates (Gentile Fusillo et al. 2019).

The first large-sky survey conducted at the Observatorio Astrofísico de Javalambre, the Javalambre-Photometric Local Universe Survey (J-PLUS) is an ongoing 12-band photometric optical survey, observing 8500 square degrees in the Northern equatorial hemisphere (Cenarro et al. 2019) at the end of the project. The filter system of this survey is a combination of broad-, medium-, and narrow-band passbands, designed to extract spectral features as the 370-400 nm Balmer break region, H_δ, Ca H+K, the G band, and the Mg *b* and Ca triplets, that are key to characterize stellar types (Jordi et al. 2006). The filters set allow to construct low-resolution photospectra for each pixel of the observed sky.

The large amount of information to be analyzed will require new and powerful data analysis techniques. In the present era, many new techniques have appeared, such as automated artificial algorithms based on machine learning techniques, that have already been used in many astrophysics studies, as the classification of galaxies (Naim et al. 1995), the identification of the counterparts of submillimetre galaxies (Liu et al. 2019), the selection of pulsar candidates (Yonemaru et al. 2018), the classification of the stars in the galactic center (Plewa 2018), the study of the galactic chemical evolution (Whitten et al. 2018) or the identification of the thin disc, thick disc, and halo WD population (Torres et al. 2019).

After realizing of the power of machine learning algorithms for classification purposes, the Random Forest classifier, a supervised classification algorithm, is used providing a classification tool for the J-PLUS survey. In preparation for the final J-PLUS catalogue, the the first data release of the J-PLUS project (J-PLUS DR1) is used as a testbed, in which the present WD population is identified.

This work is structured as follows: In Section 2 the J-PLUS filter is described. The classification method is described in Section 3. In Section 4 the first classification attempts are described. The construction of the final classification model and the comparison of its performance with previous attempts is described in Section 5. In Section 6, the resulting catalogue and some of its properties are described. In Section 7 a comparison of the catalogue from this work and the *Gaia* WDs catalogue from Gentile Fusillo et al. 2019 is made. Finally, in Section 8 the work is summarized, and some final conclusions are presented.

2. The Javalambre-Photometric Local Universe Survey

J-PLUS provides multi-filter measurements of a significant fraction of the Northern sky (scanning 8500 square degrees), by measuring the flux in 12 different photometric optical bands covering the optical range (350-1000 nm).The J-PLUS project was first conceived as a tool to support and complement the Javalambre Physics of the Accelerating Universe Astrophysical Survey (J-PAS), in the performance of the calibration tasks. J-PAS main goal is the measurement of radial baryonic acoustic oscillations using a set of 54 contiguous narrow-band optical filters.

Table 1. Description of the filter set in J-PLUS, including in last column the amount of objects that do not have measurements in each filter.

Filter	Central Wavelenght [nm]	FWHM [nm]	Missing for (of 13 405 433)
u	348.5	50.8	3 920 897
J0378	378.5	16.8	3 797 113
J0395	395.0	10.0	3 671 551
J0410	410.0	20.0	2 999 789
J0430	430.0	20.0	2 731 725
g	480,3	140.9	878 520
J0515	515.0	20.0	1 567 448
r	625.4	138.8	289
J0660	660.0	13.8	620 987
i	766.8	153.5	532 635
J0861	861.0	40.0	698 130
z	911.4	140.9	6539 34

This goal of performing the calibration of J-PAS and extracting the relevant astrophysical information from the observed sources, lead to the definition of the J-PLUS photometric passbands. A brief description of the filter set can be seen in Table 1, including the central wavelength and the FWHM of each filter. The low frequency continuum is covered by 4 SDSS filters (griz filters), joined to 6 intermediate band filters covering the blue side of the 370-400 nm Balmer break region (u filter), Ca H+K (J0395), H_δ (J0410) and the Ca triplet (J0861); and 2 narrow band filters, J0378 and J0660, sensitive to the [OII]/λ372.7 and Hα/λ656,3 emission lines (Jordi et al. 2006).

Some machine learning algorithms are not able to work with datasets with empty cells, so that only those objects with measurements in every filter can be used. In the third column in Table 1, the amount of objects that do not have measurement in each filter is described.

In April 2018, the J-PLUS DR1 was made public with measurements for more than 13M objects, allowing for Milky Way science and stellar populations studies, as for extragalactic science. J-PLUS data has already been used in different scientific goals, such as the discovery of ultracool WDs (Solano et al. 2019), the determination of star formation quenching times-scales of green valley galaxies (Nogueira-Cavalcante et al. 2019), the calculation of stellar astrophysical parameters (Carrasco et al. 2019), and the identification of low metallicity stars using artificial neural networks (Whitten et al. 2019).

3. The Random Forest Classifier

The classification method defined here to identify WDs in J-PLUS DR1 is based on the random forest classifier, a type of machine learning algorithm. Machine learning can be understood as the ensemble of methods through which the computer system is able to recognize patterns based on data analysis and then apply them for understanding new data (Mitchell 1997; Bishop 2006). There are many different applications of machine learning algorithms, including classification and data mining for big data.

One type of machine learning algorithms are the ensemble methods, those in which a finite set (an ensemble) of alternative models or classifiers is used to get a final result by combining them. Random forest is an ensemble method very useful for classification (splitting a set of objects in different classes, discrete

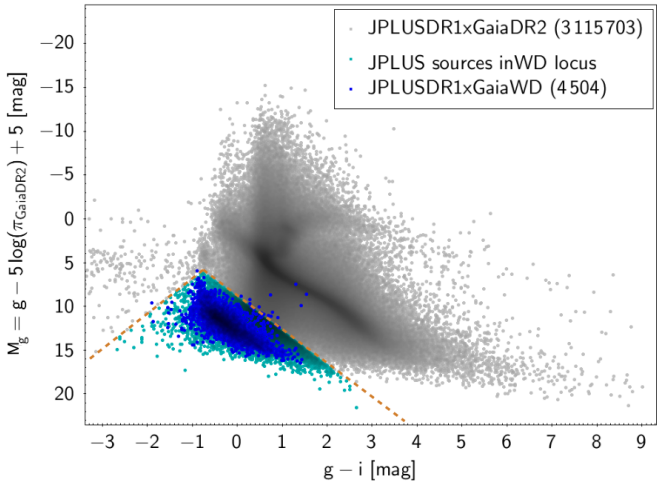

Fig. 1. Colour-magnitude diagram M_g vs (g-i). In grey, sources in J-PLUS DR1 and *Gaia*-DR2 with no missing measurements (see Table 1). WDs from Gentile Fusillo et al. 2019 present in J-PLUS DR1 used to define WDs locus in dark blue. In light blue, sources inside the defined WD locus.

values) and regression (predicting continuous values) processes. It is based in the construction of a finite set of decision trees and the combination of them to provide the final solution (Breiman 2001). Each decision tree is constructed by evaluating a data set (the training set) with a given list of attributes or characteristics (magnitudes, parallax, and their uncertainties in the current case) of each object (Quinlan 1986). This training data set includes the class (WD and non-WD in the current case) to which each object belongs, and in which the algorithm will classify other objects (the application set) without that information once it is trained. Once the training process is finished and the classification model is constructed, another set, also with complete information (attributes and classes), is used to test the model and its performance (the test data set). This procedure will provide some information about the quality of the obtained classification when working with new samples from which the class is previously unknown.

Each decision tree of the final classifier will have a root (the starting node) and a given number of branches, nodes and leafs. The classification starts from the root of the tree where the set is divided in subgroups depending on their attribute values. Each object follows the branch corresponding to its value and jump to another node, where another attribute is evaluated. This splitting continues in all the internal nodes of the tree until reaching a leaf node, where the class is predicted. Different classification trees in the forest are constructed with different roots and different analyzed attributes in each node. The final prediction is derived from the combination of all the different trees built in the forest, and a probability of getting a prediction is obtained from the mean of its probability in each tree.

The hyper-parameters of the classifier allow the user to modify its performance, tuning accordingly the quantity of leaves, branches, trees, etc. There are many parameters that the user can choose in order to characterize the model, but only few combinations of them will be optimal depending on the training set.

4. Different classification studies

Different options of increasing complexity were tried in order to get the best performance of the classification model for the J-PLUS sources. The first step was to investigate which information could be obtained from color-color diagrams using only J-PLUS magnitudes (see Sec. 4.1). Since the possible classification using color-color magnitudes was not clean, the next step (see Sec. 4.2) was to use the information of the magnitudes and their uncertainties in a classification algorithm, in order to see if the algorithm was able to recognize faint WDs because their uncertainties are bigger. The model resulted not able to disentangle WDs in the range 0<(g-i)<1 and some candidates located in and above the main sequence (MS). The third exercise (see Sec. 4.2) was to study if the classification got better using colors combining all magnitudes, instead of magnitudes individually. This study using colours is expected to provide important information about which where the best combinations of colours in order to identify WDs in a color-color diagram. The next exercise (see Sec. 5) was to include parallaxes in the classification model. The obtained results suggest in order to improve the results obtained suggesting that parallax would give a tool to classify WDs. The last exercise is to use only the most important passbands, from the importances obtained in the previous model, the bluest (u, J0378, J0395) and the reddest (i, J0861, z) ones, but no improvements are achieved.

Table 2. Summary of available sources to use in the construction of the classification model

Total number of objects in J-PLUS	13 405 433
Objects with complete information	4 595 271
Sources in *Gaia*-DR2	3 115 703
Objects outside WDs locus with Eqs. 1-2	3 081 104
Objects with *Gaia*-DR2 parallax	2 889 291
Objects with *Gaia* magnitudes (G, G_{BP}, G_{RP})	2 887 016

For all the classifications models mentioned above, training and test data sets are needed. For building these data sets, WDs and non-WDs were selected using the same procedure. WD candidates were selected from the catalogue by Gentile Fusillo et al. 2019. Obviously, only those WDs also present in J-PLUS DR1 catalogue can be used. The algorithm cannot work with a data frame containing gaps (empty cells), and so every object need to have information in every used filter. Among the WDs present in both the Gentile-Fusillo catalogue and J-PLUS DR1, only those with the probability in Gentile-Fusillo of being a WD larger than 75% (Pwd>0.75) are selected, getting a sample of 4503 identified candidates usable for training and testing purposes. From these available WDs, 2300 are used in the training and the rest are used in a test in order to validate the performance with new sources. A second test set is built using all available WDs.

The non-WD objects used to build the training and test data sets are selected from the M_g. vs (g-i) HR diagram, by excluding the WD locus defined with Eqs. 1 and 2, as can be seen in Fig. 1. This is a set of 2 887 016 available non-WD for the training and testing of the model with complete information (see table 2).

$$M_g = g - 5\log(\pi_{GDR2}) + 5 > 3.9(g-i) + 9 \quad (1)$$

$$M_g = g - 5\log(\pi_{GDR2}) + 5 > -4(g-i) + 3 \quad (2)$$

Finally, since the final classification will depend on the initial conditions in which the model has been built, it is important to avoid biases. In this sense the non-WD sources to form the training and test sets are selected by maintaining their distribution in g-i color, (Fig. 2) in order to faithfully reproduce the sample in which the final classification model will be applied. In this way we avoid selecting only sources with 0.4<(g-i)<2.8 by random

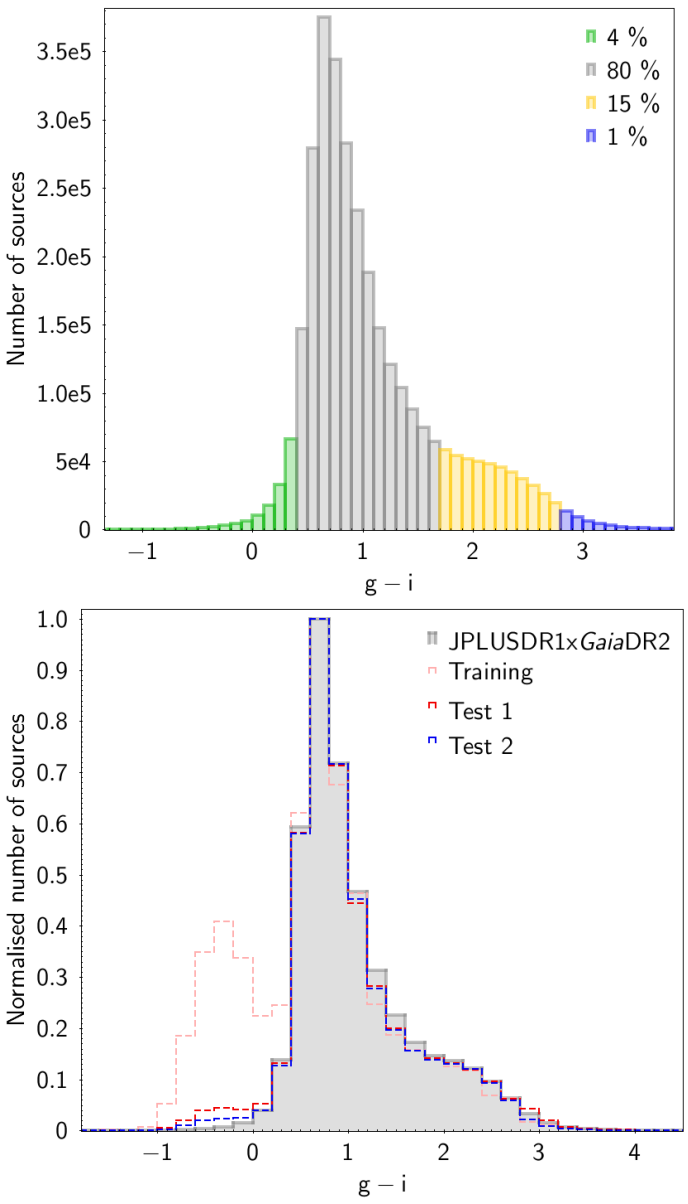

Fig. 2. *Top*: Histogram of color $g - i$ for the set of non-WDs available for the model training, showing the percentages that will be maintained in order to avoid having only sources with 0.4<g-i<2.8. *Bottom*: Histogram comparing all datasets used in this work. As these datasets contain also WDs (only non-WDs in top panel), in this histogram their contribution can be seen in -1<(g-i)<0.5. The decreasing proportion of WDs in each dataset can also be appreciated.

selection, since this is the most abundant color. As can be seen in Fig. 2, since the observation are only outside of the galactic plane, there are few blue stars, which will allow to better distinguish WDs.

As can be seen in Fig. 3, sources with all magnitudes (g passband used as an example) have been used for the training and the validation of the model, avoiding preparing the model with a distribution in magnitudes different than the one in the entire J-PLUS DR1 data set.

Under these conditions, 5000 non-WDs are selected (together with the 2 300 known WDs) to form the training set. Other 50 000 non-WDs are selected (together with the other 2204 WDs) to form the first test data set. The second test data set is built using 200 000 non-WDs and all the available WDs

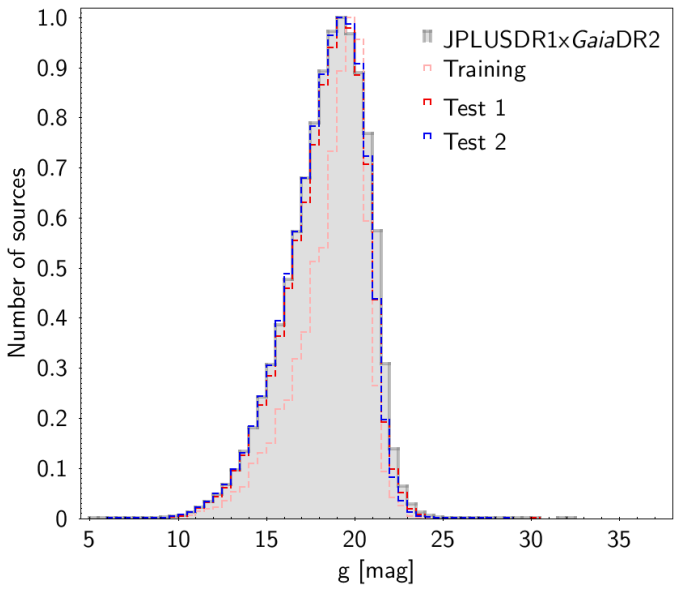

Fig. 3. Histogram of g magnitude comparing the J-PLUS data set for classification with the training, and tests sets.

(4504). This second test is used to prove that the classification model is able to recognize WDs even when they are in a smaller portion (smaller than the ∼ 25% used in the training) of the classified data set, as will happen when classifying J-PLUS data.

For every classification model, a search of the best hyper-parameters is done by applying the k-fold cross-validation (CV) technique. In a k-fold CV, the available training data is splitted into k groups and the algorithm is runned k times using $k - 1$ groups as training set and one group to validate the performance of the model. This is done in turn until each group has been left out once (Baumann 2003). In the search of the mos suitable hyper-parameters, the k-fold CV is runned for every possible model (with different combinations of the hyper-parameters). The most suitable set of hyper-parameter, and then the best model, is found when the best score (calculated for the $k - 1$ validation groups) is obtained.

Once the model is trained, making a test is a first attempt to classify new objects not present in the set with which the model has been build and evaluate its efficiency. To make a test the algorithm classifies new objects and then compares the prediction with the known value. For this purpose, once the model is built, it is validated with two tests. The first test is used to make a first validation of the model. The second test is made in order to make sure that the performance of the model does not get worse when it classifies a data set with an smaller proportion of WDs, and then it is made using all the available WDs from Gentile-Fusillo (4504 WDs) and 200 000 non-WDs. Since the aim is to classify a set of ∼3M objects that will contain a smaller proportion of WDs than the one in the training set, a second test including a bigger amount of non-WDs can help to check if the classification model is biased to find a given percentage of WDs or if it is more difficult to identify WDs when they are in a smaller proportion.

From each test, many outputs are obtained, including the accuracy of the model achieved, the confusion matrix and a classification report in which different scores of the model are evaluated. The accuracy provides the most summarized evaluation of the classification model. The confusion matrix evaluates the amount of false positives and false negatives. By comparing the predicted values (columns of the confusion matrix) with the true values (rows of the confusion matrix) given for the test, the

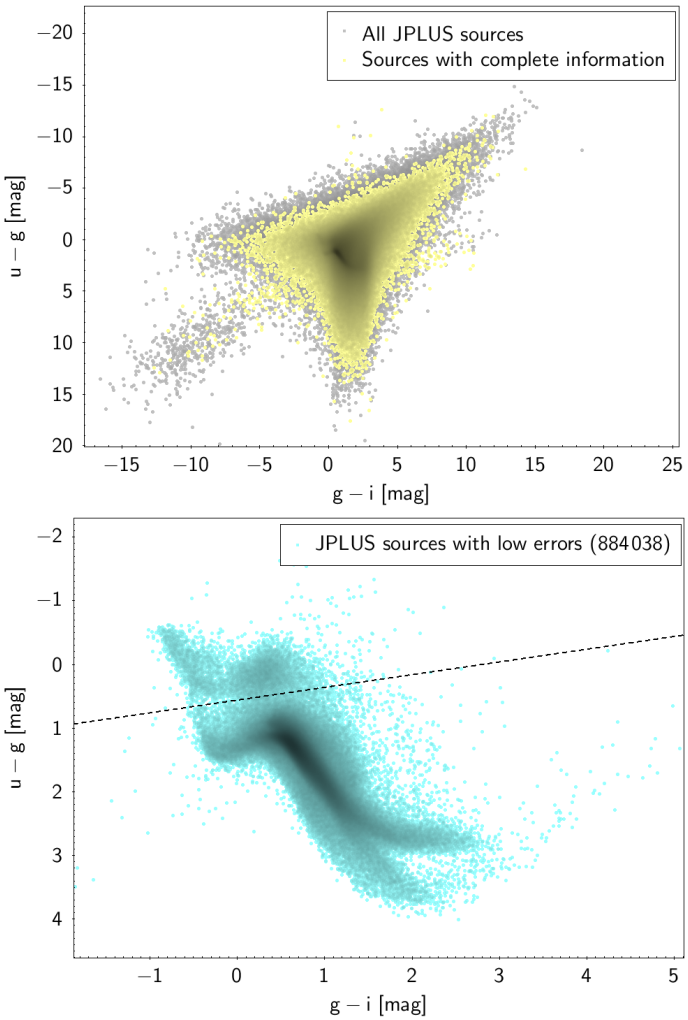

Fig. 4. u-g vs g-i diagrams. *Top*: In grey, all J-PLUS sources (13 405 433), in light yellow only those with measurement from every filter (4 595 271). *Bottom*: sources with σ_u, σ_g, σ_i <0.1 (884 038). Above the dashed line, (u-g)>-0.2*(g-i)+0.56, WDs and QSOs are located, so that QSOs and WDs models would be necessary in order to differentiate them.

amount of correctly classified objects can be seen in the diagonal of the matrix. Hence, the ideal classification matrix is diagonal, meaning that all true values are correctly predicted.

In next subsections, the first exercises in order to find the best classification model are described. The final classification model is described in Sec. 5.

4.1. Color-color diagrams

A first exercise was to make the classification using only the available J-PLUS information. We started by looking for color-color diagrams in which the WD population could be distinguished. As can be seen in Fig. 3, to obtain a good or understandable color-color diagram, it is not enough with getting only those sources with complete information, but it is necessary to filter those with large errors that somehow blur the true diagram we need to see.

In the bottom panel in Fig. 3 the same color-color diagram than in the top panel is shown, but this time using only sources with low errors (σ_u, σ_u, σ_u < 0.1). Using WD models we could try to identify the WD locus, having the majority above the

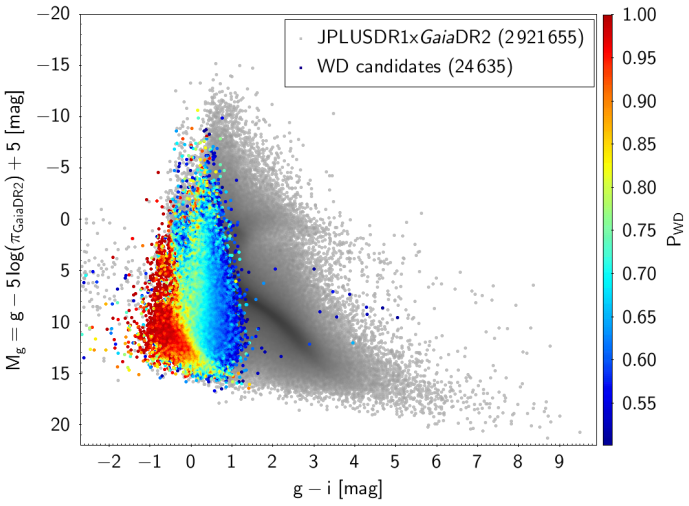

Fig. 5. HR diagram from the classification model using only magnitudes and their uncertainties. In grey, classified sources. In colours, sources with probability>0.5, considered WD candidates.

dashed line drawn in the bottom panel diagram. But cleaning up the sample using only sources with low errors leads to a smaller sample: in the bottom panel of Fig. 3, less than 1M sources (from the entire +13M in J-PLUS DR1) are shown. Using only magnitudes and color-color diagrams would allow to properly classify only few stars, as their uncertainties create confusion. Then, a classification using machine learning algorithms is done incorporating these uncertainties in the training process.

4.2. Classification using only magnitudes and their uncertainties

The first classification model using machine learning algorithm is built using a set formed by all sources present in J-PLUS DR1 with complete information in the 12 filters (4 595 271 objects). As explained in the previous section, the first step in constructing a classification algorithm begins with its training. Since this first goal is to build a model that identifies WDs using only magnitudes from J-PLUS, during the training the information of the parallax (as well as the *Gaia* filters) is not used, although it is available.

In this case, the hyper-parameters found to be the most adequate are 568 decision trees ($n\_estimators$ = 568) and 50 levels as maximum depth ($maximum\_depth$ = 50).

The confusion matrix for the two different tests (the first using 2204 WDs and 50 000 non-WDs to make a first validation, and the second test using 4504 WDs and 200 000 non-WDs to check that the performance do not get worse when the portion of WDs is smaller) is shown in Eq. 3.

$$\begin{array}{cc} \text{Test 1} & \text{Test 2} \\ \begin{pmatrix} 48\,770 & 1230 \\ 59 & 2145 \end{pmatrix} & \begin{pmatrix} 195\,015 & 4985 \\ 132 & 4372 \end{pmatrix} \end{array} \quad (3)$$

As can be seen, for the first test 2145 from 2204 WDs are recovered, obtaining a ∼ 97.3 % of success. For non-WDs, only 1230 from 50 000, are wrongly classified, with a 97,54% of success. In the second test, the 97,5% of non-WDs and the 97.07 % are correctly classified. This results seem to show that changing the percentage of WDs in the data set does not make the

performance of the model get worse, since the amount of well classified objects does not change too much from the first to the second test, in which the amount of WDs goes from being the ∼4% (in the first test) to the ∼2% (in the second test).

In Fig. 5 WDs candidates obtained from this classification model (using only magnitudes and errors) are plotted in a color-magnitude diagram in order to judge their fiability, recovering the parallax information after the classification (parallax is not used during the classification process). An horizontal gradient in probability can be seen, showing that using only magnitudes (and their uncertainties), the classification model is not able to distinguish between objects lying in the same vertical line in the diagram. Giving the high amount of WD contaminants commented from the confusion matrices, considering the parallax information from *Gaia*-DR2 seems to be the way to improve the results.

4.3. Classification using colors

The colours of the stars should be kept independent of the brightness of the source (if the effect of the uncertainties is neglected). Then, in principle, defining a classification algorithm using colours instead of magnitudes could be a good method to disentangle the WD population from the rest of objects. Besides, in this way, the ranking of the most important attributes obtained from the training of a model could give some clues on which could be the best color-color diagrams to identify WDs. Training the model using all combinations of the 12 J-PLUS filters, all the colors, gives the possibility of getting a ranking with the most important colors and see if it provides new information.

Using again the same training and tests sets used in previous section, but having now 134 attributes, all colors, instead of 12 filters, the search of the most suitable hyper-parameters gives new results, being the maximum depth of each tree (*max_depth*) equals to 30 for 676 classification trees in the forest (*n_estimators*).

In Eq. 4 the confusion matrix for each test can be seen.

$$\begin{array}{cc} \text{Test 1} & \text{Test 2} \\ \begin{pmatrix} 49\,033 & 967 \\ 128 & 2076 \end{pmatrix} & \begin{pmatrix} 196\,225 & 3775 \\ 156 & 4348 \end{pmatrix} \end{array} \quad (4)$$

The results are not getting better from the obtained in the classification model using magnitudes instead of colors, since for both tests less WDs are correctly classified. Although more non-WDs are correctly classified, meaning a less amount of contaminants in the set of sources classified as WDs, more WDs are lost in the sources classified as WDs. Moreover, the percentages of contaminants in the WDs set (non-WDs classified as WDs) have not been improved so much.

The most useful information we can get from this classification model is the most relevant combinations of colors. The top 8 colours in the ranking of importances (from ∼7.52% to ∼4.99% of importance) are occupied by the combinations of the u filter with 4 other filters (u-i, u-J0861, u-z and J0660-u), the same 5 filters occupying the top positions in the ranking of the model using only magnitudes. Therefore, with this model no more information neither better results are obtained. This means that the algorithm takes into account the combination of the information in the attributes, and since using colors is simply a lineal combination of magnitudes, no new information is added by using colors instead of magnitudes.

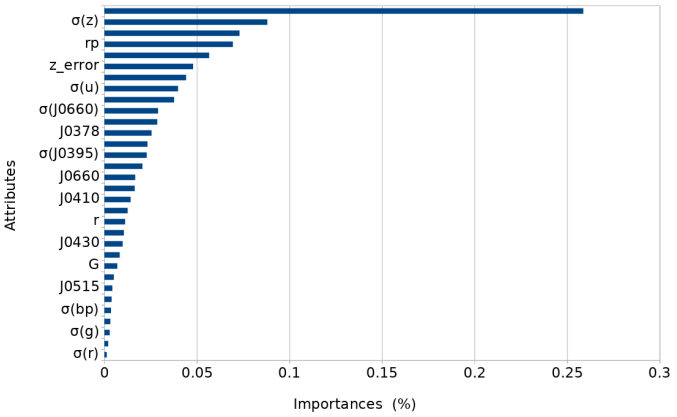

Fig. 6. Feature importances for each variable from the training of the model using magnitudes, parallax and errors. This importances show how significant is the value in each attribute to decide the class of an object.

5. Training and testing of the final model

The results obtained in Fig. 5 clearly suggested that the use of the parallax for the classification is a very important tool to better classify WDs. Since all sources in future J-PLUS DRs are expected to have a counterpart in *Gaia* measurements, and then the parallax information will be available, in this final classification model, *Gaia*-DR2 parallax and its error are included as new attributes. For this, the selection of sources to build training and test sets is made in the same way than for the previous attempts, getting 2 887 016 non-WDs and 4503 WDs with complete information in 12 J-PLUS passbands, 3 *Gaia* passbands (G, $G_{\rm BP}$ and $G_{\rm RP}$) and parallax (including their uncertainties) available for building the training and test sets (see Table 2).

5.1. Training of the classification model including parallax

For the training process, as explained before, the test set is made up of 7300 sources, including 2300 WDs selected from Gentile Fusillo et al. (2019) with complete information in all J-PLUS and *Gaia* passbands and parallaxes, and the flag indicating the class (WD or non-WD) extracted from Gentile Fusillo. Using this training set, we look for the best set of hyper-parameters of the model, obtaining 137 decision trees ($n\_estimators = 137$), 110 levels as the maximum depth of each tree ($max\_depth = 110$) and the maximum number of features to consider when looking for the best split as the square root of the number of features ($max\_features$ ='sqrt'), that in our case is $\sqrt{32}$ (12 J-PLUS filters, 3 *Gaia* filters, parallax and errors).

Once the training, and thus the construction of the model, has been done, the feature importances can be analyzed, indicating how significant is the evaluation of each feature for the classification. As it can be seen in Fig. 6 the parallax has the predominant importance of the information in the classification (∼26%). Without taking the parallax measurement into account, any object with $(g-i) < 2$ mag was considered to be a WD by the algorithm (see section 4.2). In this new classification model, parallax helps to distinguish between WDs and hot main sequence stars. The importance of the parallax result to be very high in comparison with the rest of the information, but looking at Fig. 6 it can be seen that filters z, J0861 and G_{RP} are the most important passbands, including their uncertainties. These are the reddest passbands, and just after them in the ranking we find the

Table 3. Classification report from the first test of the final model (using magnitudes, parallax and errors). In each column, the evaluated precision, recall and f1-scores (see definitions in Sec. 5.2) for each class (0=non-WDs, 1=WDs) are shown.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	50 000
1	0.97	1.00	0.98	2204
avg / total	1.00	1.00	1.00	52 204

Table 4. Classification report from the second test of the final model (using magnitudes, parallax and errors) In each column, the evaluated precision, recall and f1-scores (explained in Sec.4) for each class (0=non-WDs, 1=WDs) are shown.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	200 000
1	0.94	1.00	0.97	4504
avg / total	1.00	1.00	1.00	204 504

bluest ones, u and J0378. The intermediate passbands seem to play a secondary role in the classification process, since these 5 passbands and their uncertainties sum up the ∼49% of importance. The importances of these 5 filter added to the importance of the parallax and its errors gets the ∼77.3% while the other 10 passbands and their uncertainties get very low importances, with only the ∼23.7% of importance shared among all.

5.2. First test of the model using magnitudes, parallax and uncertainties

For this test 50 000 non-WDs are selected in the same way that for the training set (by random selection but maintaining the original distribution in color, see Sec.4). The WDs not selected for the training set are used now, that is 2204 WDs from Gentile Fusillo et al. (2019) present in J-PLUS DR1.

As explained before, the ideal confusion matrix would have only values in its diagonal and zeros otherwise, and in this case it would be a diagonal 2x2 matrix with values 50 000 and 2204. Instead, for the first test of the model, the resulting confusion matrix is the following:

$$\text{Test 1} \qquad \begin{pmatrix} 49\,936 & 64 \\ 8 & 2196 \end{pmatrix} \tag{5}$$

In the Eq. 5 containing the confusion matrix it can be seen that 64 over 50 000 non-WD (∼0.13%) and 8 over 2204 WD (∼0.36%) have been wrongly classified by the classification method. The results of this first test are much better than the results obtained in the previous classification models.

As explained is Sec. 4, in the classification report in table 3 the information of the calculated *precision*, *recall* and *f1-scores* for each of the different classes (two in this case) it is shown. The *precision*, also called the positive predictive value, is calculated as the number of correct predictions divided by the number of correct predictions plus false positives, being 2 196/(2 196+64)≃0.9717 for WDs 49 936/(49 936+8)≃0.9998 for non-WDs. The *recall* is the number of correct predictions divided by the total number of instances, so $tp/(tp + fn)$ being 2196/(2196+8)≃0.9964 for WDs 49 936/(49 936+64)≃0.9987 for non-WDs. The *f1-score* is defined as the weighted harmonic mean of both the precision and recall, having that the best f1-score is 1 and the worst value is 0. As can be seen in table 3, the averaged *f1-score* is 1, that is the best possible result. Finally, the *support* is simply the number of instances that are the correct target values.

As is shown, the results of this test of the model are very good. The *accuracy* score calculated for the results of the test gives the probability of getting the good classification for each object of the sample. In the current test, the accuracy score obtained is ∼99.86% indicating a very good performance of the model in classifying the objects present in the sample.

5.3. Second test of the model using magnitudes, parallax and errors

For the second test of the model using magnitudes and parallaxes, 200 000 non-WD (different from those selected used in the previous test) are selected and all the WDs from Gentile Fusillo et al. (2019) present in J-PLUS DR1 with complete information, that is 4504 (2300 already used for the training and the other 2204 used in the first test), as explained in Sec. 4. The obtained confusion matrix is shown below:

$$\text{Test 2} \qquad \begin{pmatrix} 199\,732 & 268 \\ 12 & 4492 \end{pmatrix} \tag{6}$$

In the Eq. 6 containing the confusion matrix of this second test, it can be seen that 268 over 200 000 non-WD ∼0.13% has been wrongly classified, being almost the same percentage than the one obtained in the first test of this model, ∼0.134%. For the WDs, only 12 over 4504 WD (∼0.27%) have been wrongly classified by the model, that is only 4 more errors than in the first test but having now twice as many WDs as before (4504 WDs in test 2, in front of 2204 WDs in test 1). Again, the results are not getting worse for the second test, in which the percentage of WDs is half the percentage in the first test.

In table 4 we can see the classification report for this test. Again, very good results are obtained even though the amount of non-WDs against the amount of WDs is twice the fraction used for the first test. All the calculated scores (*precission*, *recall* and *f1-score*) get an averaged value of 1, the best possible result. Moreover, the accuracy score obtained is again a ∼99.86% showing that we have a reliable model to use in the identification of the WD population present in J-PLUS DR1.

The ranking of importances obtained with this last classification could be used as a tool to select the most important passbands in order to build another model using only those. Having less passbands could make the algorithm faster, and enable to classify a bigger portion of objects, adding those that do not have gaps in the most important passbands (perhaps removed in this classification because they do not have information in passbands that we would not use in a possible new model). Classifying using only the most important passbands, makes the accuracies worse and do not add many objects to the classification. Even leaving out passbands including J0660, we don not even get more than 88 000 new (from those available for the classification using all passbands and their uncertainties, 2 921 655 sources) available objects (from the total of 13 405 433 in J-PLUS DR1),

Table 5. Comparison between 3 different classification models (1=using only magnitudes, 2=using colors, 3=including parallaxes). In the first column it is shown the accuracy of the model, in the second column the percentages of non-WDs wrongly classified (FN=false negatives) are included, and in the third column the percentages of WDs wrongly classified for the first and second tests (FP=false positives).

model	accuracy %	FN %		FP %	
1	~97.5	2.68	2.93	20.2	24.9
2	~98	5.81	3.46	19.3	18.9
3	~99.9	0.36	0.27	0.13	0.13

and the accuracies get worse. Moreover, the built classification model using all magnitudes, parallaxes and uncertainties is anyway very fast in the classification. Then the improvement of the CPU time to perform the classification is not yet an issue at this stage and the loss in performance when removing information made this option less optimal and not needed.

5.4. Comparison of all classification models

Finally, the results obtained in the three different attempts of classifying with machine learning algorithms are summarized in Table 5. In the first column of the Table 5 it can be seen how the accuracies have been improved in each exercise, until obtaining the best one for the model that includes the parallax information. In the second column of the Table 5 the percentages of non-WDs wrongly classified as WDs (FN) is shown, also improved in the last model. Finally, the best improvement obtained can be seen in the third column, where the percentages of WDs wrongly classified as non-WDs are shown. Using parallaxes in the classification model makes its performance much better, as can be seen from all the results.

6. Final WD candidates

Once the results from the test make it possible to feel confident about the performance of the classification model, the next step is the classification of the sample of J-PLUS DR1 sources. The most important constrain in order to discard a source to be classified, as has been commented during this work, is a missing measurement from a passbands, since the classification algorithm uses all magnitudes and is not able to classify an object with missing information. Only 4 595 271 objects (~34.3% of the catalogue) in J-PLUS have complete information in the 12 J-PLUS passbands (table 2). It can be assumed that future releases will be more complete and then it is preferable to continue classifying using all the passbands instead of taking less passbands and then classify more objects. As explained before, for the building of the final classification model, parallaxes from *Gaia*-DR2 were used, so that only sources with this information available can be classified using that model. There are 3 115 703 objects in both J-PLUS DR1 and *Gaia*-DR2 catalogues, with complete information in the J-PLUS filters, but when applying filters to get only sources also with complete information from *Gaia* -DR2, we end up with 2 921 655, that is a ~21.8% of the entire catalogue from the first data release.

When the classification model defined in Sec. 5 is applied to the sample of 2 921 655 sources from J-PLUS DR1, the classification model returns the probability of each object of being classified as a WD. In Gentile Fusillo et al. 2019, a probability of being a WD is calculated for every WD candidate. In contrast, the probability obtained from the classification model in this work, is not exactly the probability of being a WD but the probability of being classified by the model as a WD. It basically gives an assessment of the reliability of the classification and it is not a typical probability since it is not driven by any PDF (Probability Distribution Function). In principle, with an accuracy above the 99% for two different tests, we can rely on the results and the 50% threshold in "probability", but lets visualize these results.

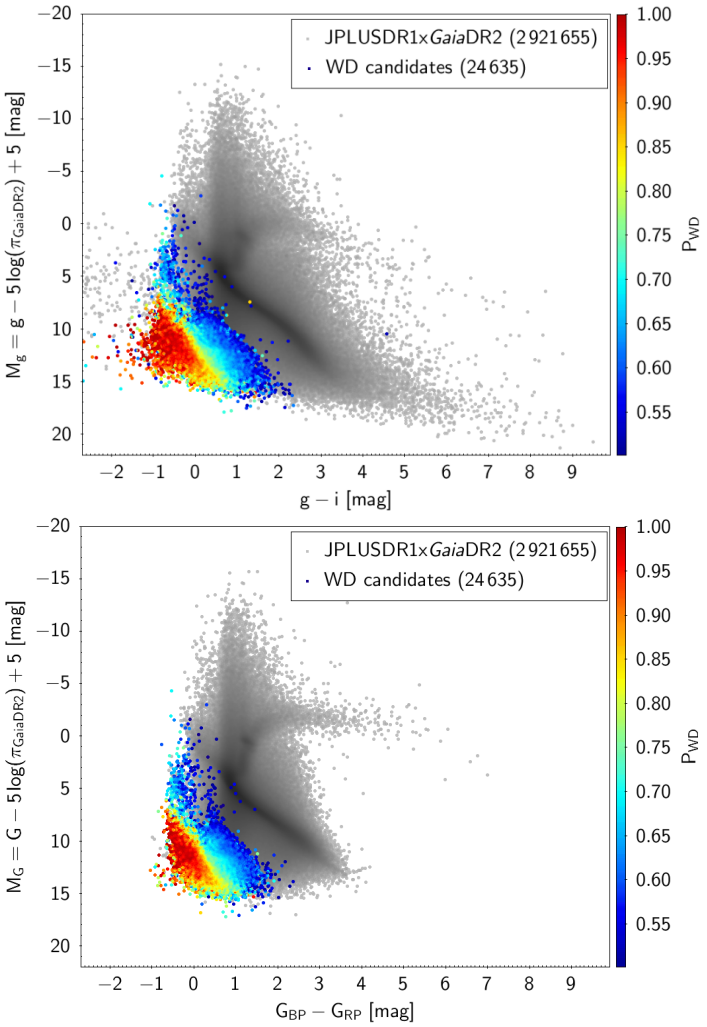

Fig. 7. Color-magnitude diagrams from the classification model using magnitudes, parallax and errors. In grey, classified sources. In colours, sources with probability>0.5, considered WD candidates. *Top*: Diagram with J-PLUS filters. *Bottom*: Diagram with *Gaia* filters.

Taking as WD candidates the objects with a probability from the model higher than 50%, 24 634 WD candidates are obtained. From the results obtaining when testing the model (in Sec. 5), the amount of contaminants in the set of WD candidates can be inferred. Since for both tests, only ~0.13% of the non-WDs where wrongly classified as WDs, it can be expected that from the sources classified at the end (2 921 655 sources) there will be a set of ~4000 sources wrongly classified as WDs, constituting a set of contaminants.

The color-magnitude diagram for the set of classified sources can be seen in Fig. 7 where we can clearly see that the vast majority of the candidates are located in the WD locus. We can see that the objects with the highest P_{WD} are confined between -1 and +1 mag in both diagrams. In fact this is where the WD locus is located. Moreover, the probability of being classified

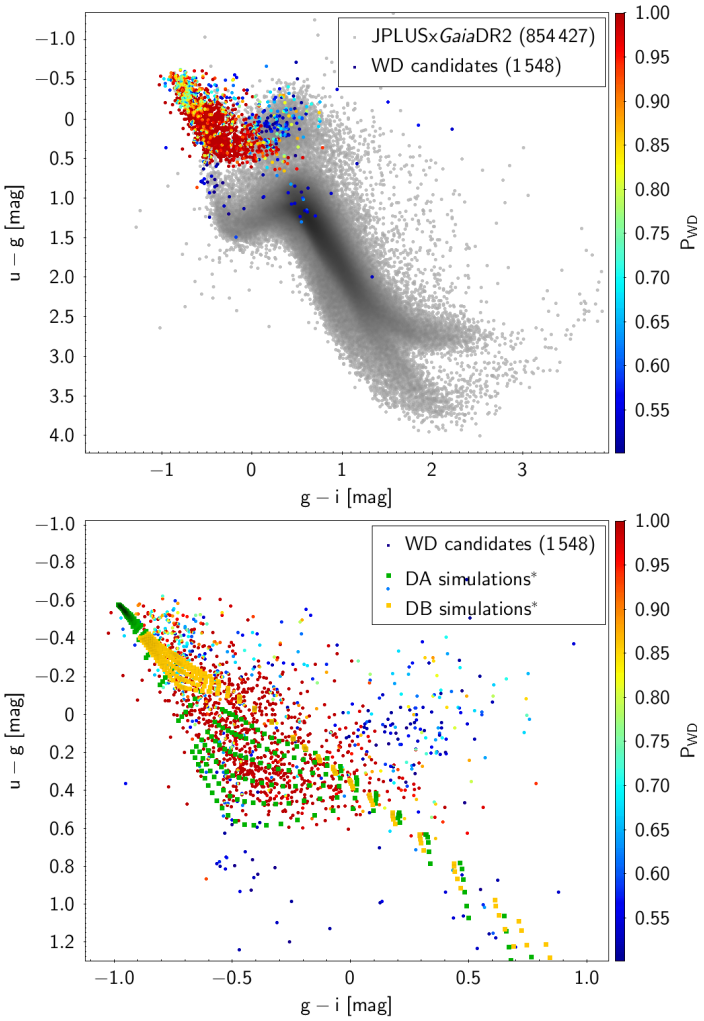

Fig. 8. u-g vs g-i diagrams. *Top*: In grey, classified J-PLUS sources with low errors, in colors WD candidates. *Bottom*: Zoom to the WDs locus and models from http://www.astro.umontreal.ca/ bergeron/CoolingModels added.

as a WD seems to decrease with the distance to the WD locus. We also find some candidates in the location of the main sequence, but fortunately their associated probability is below 55%, and they can be identified as the contaminants we expected from the classification commented in Sec. 4. These possible contaminants can be studied by taking the objects with $M_g = g - 5\log(\pi_{GDR2}) + 5 < 3.9(g-i) + 9$, obtaining 3373 with probabilities between 50% and 98%. From these 3373 sources, there are 251 WDs candidates with a probability of being classified as a WD higher than 75%, that are be wrong located in the diagram because of their big uncertainties. Although it could seem that fixing the threshold in probability at 60% would give more reliable results, there are also objects located in the WD locus with low errors but a probability below this threshold. This is due to the fact that the probability we are talking about cannot be directly treated as "a probability of being a WD", as commented before. That is to say, that an object can be well known as a WD and still have P_{WD}<60% because for some reason it is 'difficult' to classify it as a WD. The entire sample of WD candidates is presented, to be used taking into account that the model has a ∼99% of accuracy and the result still contains some contaminants that can be treated as wanted using the information of the probability.

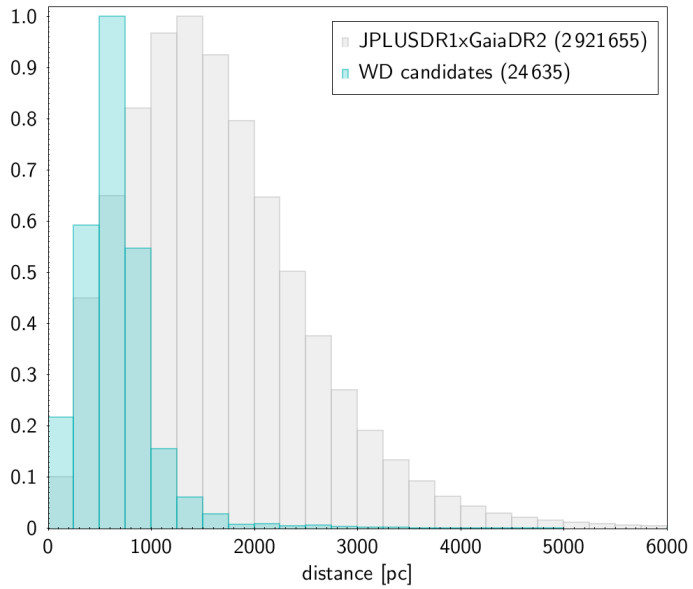

Fig. 9. Sources vs distances from Bailer-Jones et al. 2018. In blue, WD candidates. In grey, all the classified sources. The bins are normalized to the maximum in order to visualize both distributions, since the amount of WDs is much smaller.

Figure 8 shows a color-color diagram (u-g)vs(g-i) with all sources with σ_u, σ_g and σ_r <0.1 mag. WD candidates are shown in the top panel with coloured points, with a gradient in the probability of being classified as a WD. In the bottom panel, the models for DA and DB subclasses (Holberg, & Bergeron 2006; Kowalski, & Saumon 2006; Tremblay et al. 2011; Bergeron et al. 2011) are shown. The farther classification of WDs candidates found in this work in WD subclasses is left to future work. It can be seen that WD candidates with higher probability are located in the correct locus, while the candidates that are not near any of the two models lines have probabilities below 55%. Candidates in (u-g)<0.5 with low probabilities (P_{WD} <0.55) that are not near any of the two simulations (DA and DB simulations) could be sources confused with QSOs since, as commented in Sec. 4.1 for Fig. 3.

The amount of objects in function of the distance derived in Bailer-Jones et al. 2018 is shown in Fig. 9. The bins in this figure are normalized in order to have the maximum of each distribution to 1, since the amount of WDs candidates is very small compared with the entire set (24 635 of 2 921 655). Although the entire data set contains sources beyond 13 000 pc, WDs candidates are no longer present after 7000 pc. It is completely understandable, since the farthest WDs can be too faint to be measured.

7. The *Gaia* DR2 catalogue of white dwarfs

In this section the sources classified as WDs in this work will be compared to those in Gentile Fusillo et al. 2019. From the crossing of this WD catalogue with J-PLUS DR1 and the selection of sources with measurements from all passbands and a probability higher than 75% ($P_{WD} > 0.75$) 4504 WDs were obtained. From cross-matching WD candidates in this work with the 4504 selected WDs from Gentile Fusillo et al. 2019, a total of 4488 sources in common are recovered. In Fig. 10 a comparison between the probability provided by Gentile Fusillo et al. 2019, and the probability derived from the classification model in this work is shown. From the 19 WDs that are not recovered by the classification model, 3 are not in the classified set (because they

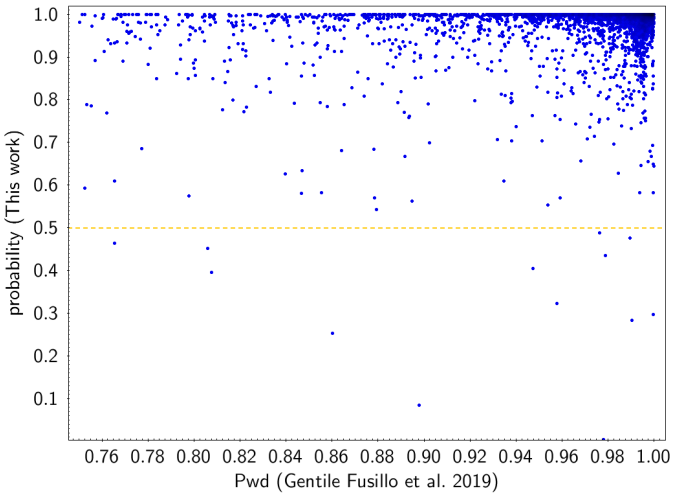

Fig. 10. Probability (from this work) vs Pwd (from Gentile Fusillo et al. 2019). The dashed line in yellow separates candidates (below the line) from sources that are considered high confidence candidates by Gentile Fusillo et al. 2019 but they turn out not to be candidates after the classification in this work.

Table 6. Sources in the entire set from Gentile Fusillo et al. 2019 (486 641) and the classified set (2 92155). From this cross, 7128 sources are found: 6143 are candidates from this work, while 985 sources have probability<0.5 and then are not classified as WD candidates (First and second columns.

Prob<0.5	985	Pwd>0.75	12
		Pwd<0.75	973
Prob>0.5	6143	Pwd>0.75	4488
		Pwd<0.75	1655

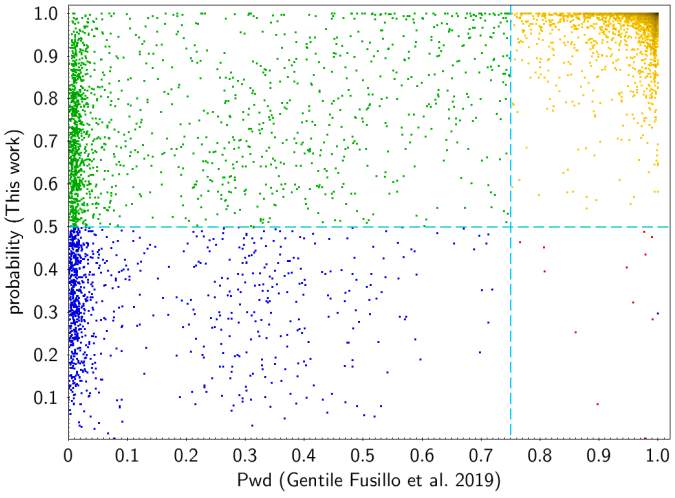

Fig. 11. Probability (from this work) vs P_{wd} (from Gentile Fusillo et al. 2019 for the 7128 obtained when crossing the entire Gentile-Fusillo (486 641 WD candidates) catalogue with the entire classified data set in this work (2 921 655 sources). Dashed lines in light blue separate groups in table 6.

do not have measurements from some of the passbands or parallax), and 16 obtain a probability from the model below 50%, and so not getting to be a candidate. These sources can be seen in Fig. 10, having them below the dashed line.

It can be seen (in Fig. 10) that from the 16 sources that are high confident WDs candidates for Gentile-Fusillo but do not get to be a candidate in this work, there are 1 source that was not

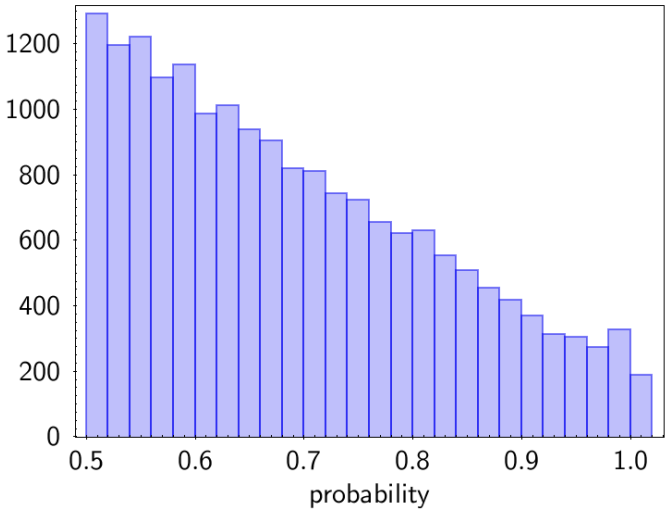

Fig. 12. Probability distribution of WDs candidates not present in Gentile-Fusillo

classified as a WD in any tree in the forest of the classification model, since it is obtaining a probability below 10%. The rest of objects could be the errors obtained from the classification model, since although working with accuracies higher than the 99% some wrong classifications can always be expected when working with machine learning algorithms, as seen in Sec. 5.

When crossing the entire classified set (2 921 655 sources) with the entire WDs catalogue from *Gaia* DR2 (486 641), 7 128 sources are found. A summary of these sources are shown in Table 6, using "Prob" when referring to the probability of being classified as a WD obtained from this work, and "P_{WD}" when referring to the probability from Gentile Fusillo et al. 2019. From the 7128 coincident sources found, 985 are not candidates (Prob<0.5) and 6143 sources are also candidates from this work. From the non-WD candidates (985), there are 12 sources that were selected as high confidence WD candidates in Gentile Fusillo et al. 2019. The rest 973 sources (Prob<0.5 and Pwd<0.75) where not considered as high confidence WD candidates by Gentile-Fusillo, and have not been found as candidates in this work. From the 6143 WD candidates from this work (Prob>0.5), 4488 are high confidence candidates in Gentile-Fusillo while there are 1655 sources not considered as high confidence in Gentile-Fusillo, but considered candidates in this work. This results can again be better understood seeing the diagram comparing both probabilities shown in Fig. 11

If both data sets where completely compatible, in Fig. 11 would only show points in top-right and bottom-left panels. In the top-right panel, we can see in yellow sources that are high confidence WDs candidates from Gentile-Fusillo (P_{wd}>0.75) and WD candidates from the classification model in this work. In bottom-left panels there are shown in blue sources that are candidates in Gentile-Fusillo with P_{wd}<0.75, and then not considered high confidence candidates, and get also a low probability (<0.5) from the classification model in this work.

Apart from this 7128 analyzed sources found in both catalogues, in the 24 635 candidates obtained in this work there are 18 491 candidates that are not present in Gentile Fusillo et al. 2019, neither with high nor low probability. In Fig. 12 we can see the distribution in probability of these possible "new" WD candidates. Although the amount of WD candidates decreases with probability, there are still some (1776 WD candidates) sources with a probability above 90% that can be defined as "new" high

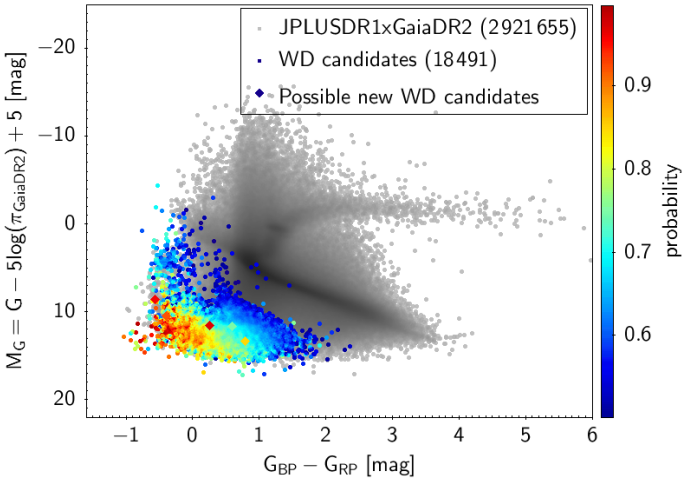

Fig. 13. HR diagram showing 18 491 WD candidates from this work that are not present in the catalogue from Gentile Fusillo et al. 2019. From these, sources that pass all quality filters by Getile-Fusillo are shown with diamons.

confidence WD candidates. This sources can be seen in Fig. 13, in the *Gaia* HR diagram. From these 18 491 WD candidates not present in the Gentile Fusillo catalogue, there are 16 128 sources with PARALLAX_OVER_ERROR < 1, and then passing the first filtering in Gentile Fusillo et al. 2019, but from them only 8816 are inside the WD locus defined in the HR diagram in Gentile Fusillo et al. 2019. Finally these 8816 sources, only 5 WD sources pass the quality filters in Gentile Fusillo et al. 2019. These 5 sources are shown with diamonds in Fig. 13, and that can be seen that their probabilities of being classified as a WD obtained in this work is higher than the 75%.

Although the classification through the HR diagram requires a larger filtering in uncertainties and quality, the classification model using machine learning algorithms have demonstrated to be a tool to make correct classifications without filtering, since it can take those uncertainties into account during the training of the model.

8. Summary and conclusions

Different ways to classify WDs using only photometry were tested in this work. Results improved when considering also parallax information to disentangle between WDs and non-WDs. The final classification method consists on a supervised Random Forest classifier, using photometry from J-PLUS and magnitudes, and parallaxes from *Gaia*-DR2. WD candidates from Gentile Fusillo et al. 2019 were used for training and validation purposes, taking only those with probability larger than 75% (P_{WD}>0.75), obtaining 4504 high-confidence WD candidates. This selection was placed in a color-magnitude diagram with sources from J-PLUS DR1 that have a counterpart in *Gaia*-DR2 and have complete information (measurements in all passbands). In this diagram the WD locus is defined in order to select those sources outside it as non-WD sources, that with WDs from Gentile Fusillo et al. 2019 will serve to form training and test sets. A classification model using magnitudes from 15 filters, parallax and errors was then built, achieving accuracies higher than 99%. All scores shown in classification reports are close or equal to 1, the best possible result. In a comparison with the previous built classification models, the one including the parallax during the classification results to be the one identifying the maximum number of objects with the minimum amount of contaminants, even when increasing the percentage of non-WDs in the test data test. These results give high confidence on the classification model, to be applied to the entire J-PLUS DR1 data set.

When applied to all sources present in J-PLUS, with complete information (2 921 655 sources) the model returns 24 635 WD candidates. For every source the model also gives a probability of being classified as a WD. The locus of the WD candidates in a color-magnitude diagram, coincides to be the WD locus for both J-PLUS and *Gaia* filters. Even though there are also some WD candidates located in both main sequences, they result to be 127 objects with probabilities between 50-76%. As there are also WD candidates with probability below 76% that are well located in many diagrams, we decide to not reduce the threshold in probability to disentangle between WDs and non-WDs, and give to possible users of the WDs candidates catalogue the possibility to decide using the information analyzed in this work.

Using simulations for DA and DB WD (from Holberg, & Bergeron 2006; Kowalski, & Saumon 2006; Tremblay et al. 2011; Bergeron et al. 2011), we find that some sources were classified as WDs and could have been confused with QSOs, since they are located in the WDs locus in the u-g vs g-i diagram. This sources are returned with a low probability of being classified as a WD (P_{WD}).

A comparison of WD candidates obtained in this work with WDs from Gentile Fusillo et al. 2019 was done, finding that 4488 from the 4504 WDs candidates used during the all process are recovered. In the rest of WDs candidates there are 1655 sources also found in Gentile-Fusillo, but with P_{WD}<0.75. Finally, 18 491 WDs candidates from this work are not present in Gentile-Fusillo, and their probabilities result to be distributed for all values (higher than 0.5), containing also candidates with 100% of probability of being classified as a WD by the classification model. In a further analysis with the filtering used in Gentile Fusillo et al. 2019 for the classification, it was found that almost all these 18 491 WDs were not classified by Gentile-Fusillo because they do not pass the quality filters. The rest 5 sources that pass the filtering in Gentile Fusillo et al. 2019, have a probability from the model in this work larger than 76 %

The method built in this work for the identification and classification of WDs, using J-PLUS and *Gaia* magnitudes, and *Gaia* parallaxes, result to be a powerful tool in the classification and analysis of future J-PLUS data releases, since it is expected that all sources in J-PLUS will have a counterpart in *Gaia*. Then, this work opens the way in the identification of more WDs, enlarging the available WDs sources samples.

Acknowledgements

This work was made using data from the Observatorio Astrofísico de Javalambre mission J-PLUS (http://www.j-plus.es/home/home) and the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia) and models from Synthetic Colors and Evolutionary Sequences of Hydrogen- and Helium-Atmosphere White Dwarfs (http://www.astro.umontreal.ca/~berg-eron/CoolingModels). I would like to thank first my tutor Dr. Josep Manel Carrasco, for trusting and guiding me during the realization of this work. I also want to thank the DPAC group, especially Roger Mor and Alfred Castro, for giving their help whenever I have needed it.

References

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., et al. 2018, AJ, 156, 58
Baumann, K. 2003, Trends in Analytical Chemistry, 22
Bergeron, P., Wesemael, F., Dufour, P., et al. 2011, ApJ, 737, 28
Bishop, C. M. 2006, Pattern Recognition and Machine Learning
Breiman L., 2001, Machine Learning, vol. 45, Kluwer Academic Publishers., Dordrecht, The Netherlands, p. 5
Campos, F., Bergeron, P., Romero, A. D., et al. 2016, MNRAS, 456, 3729
Carrasco, J. M., Jordi, C., & Masana, E. 2019, Highlights on Spanish Astrophysics X, 548
Cenarro, A. J., Moles, M., Cristóbal-Hornillos, D., et al. 2019, A&A, 622, A176
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Doherty C. L., Gil-Pons P., Siess L., Lattanzio J. C., Lau H. H. B. 2015, MNRAS, 446, 2599
Fontaine, G., Brassard, P., Bergeron, P. 2001, The Astronomical Society of the Pacific, 113, 409
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Garcia-Berro, E., Hernanz, M., Isern, J., et al. 1988, Nature, 333, 642
García-Berro, E., Torres, S., Althaus, L. G., et al. 2010, Nature, 465, 194
Gentile Fusillo, N. P., Tremblay, P.-E., Gänsicke, B. T., et al. 2019, MNRAS, 482, 4570
Hansen, B. M. S., Kalirai, J. S., Anderson, J., et al. 2013, Nature, 500, 51
Holberg, J. B., & Bergeron, P. 2006, AJ, 132, 1221
Ibeling D., Heger A. 2013 ,ApJ, 765, L43
Isern J., Artigas A., & García-Berro, E. 2013, European Physical Journal Web of Conferences, 05002
Jordi, C., Høg, E., Brown, A. G. A., et al. 2006, MNRAS, 367, 290
Kowalski, P. M., & Saumon, D. 2006, ApJ, 651, L137
Lauffer G. R., Romero A. D., Kepler S. O. 2018, MNRAS, 480, 1547
Mitchell, T. 1997, Machine Learning. McGraw Hill. p. 2.
Liu, R. H., Hill, R., Scott, D., et al. 2019, arXiv e-prints, arXiv:1901.09594
Naim, A., Lahav, O., Sodre, L., et al. 1995, MNRAS, 275, 567
Nogueira-Cavalcante, J. P., Dupke, R., Coelho, P., et al. 2019, arXiv e-prints, arXiv:1907.11244
Pelisoli, I., & Vos, J. 2019, MNRAS, 488, 2892
Plewa, P. M. 2018, MNRAS, 476, 3974
Reid, I. N. 2005, ARA&A, 43, 247
Quinlan J. 1986, Induction of decision trees, vol. 1, Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 81
Rowell, N., & Kilic, M. 2019, MNRAS, 484, 3544
Solano, E., Martín, E. L., Caballero, J. A., et al. 2019, A&A, 627, A29
Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, ApJ, 750, 99
Torres, S., Cantero, C., Rebassa-Mansergas, A., et al. 2019, MNRAS, 485, 5573
Torres, S., García-Berro, E., Althaus, L. G., et al. 2015, A&A, 581, A90
Tremblay, P.-E., Bergeron, P., & Gianninas, A. 2011, ApJ, 730, 128
Tremblay, P.-E., Cukanovaite, E., Gentile Fusillo, N. P., et al. 2019, MNRAS, 482, 5222
von Hippel, T., & Gilmore, G. 2000, AJ, 120, 1384
Whitten, D. D., Placco, V. M., Beers, T. C., et al. 2018, Rediscovering Our Galaxy, 383
Whitten, D. D., Placco, V. M., Beers, T. C., et al. 2019, A&A, 622, A182
Winget D. E., Hansen C. J., Liebert J., van Horn H. M., Fontaine G., Nather R. E., Kepler S. O., Lamb D. Q., 1987, ApJ, 315, L77
Woosley S. E., Heger A. 2015, ApJ, 810, 34
Yonemaru, N., Takahashi, K., Kumamoto, H., et al. 2018, arXiv e-prints, arXiv:1811.11478
York, D. G., Adelman, J., Anderson, J. E., et al. 2000, AJ, 120, 1579
Zhao, G., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Research in Astronomy and Astrophysics, 12, 723